Technological Forecasting & Social Change 180 (2022) 121730

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

The socio-economic value of scientific publications: The case of Earth Observation satellites

Valentina Morretta [a], Davide Vurchio [b,*], Stefano Carrazza [c]

[a] Università degli Studi di Milano, Department of Economics, Management and Quantitative Methods, Italy
[b] Università degli Studi di Bari "Aldo Moro", Department of Economics and Finance, Italy
[c] Università degli Studi di Milano, Department of Physics, Italy

ARTICLE INFO

JEL codes:
H54
D61
D80
O30
Q55

Keywords:
Earth observation
Cost-benefit analysis
Research infrastructures
Marginal Social Value
Scientific publications

ABSTRACT

This paper provides a conceptual framework to estimate the socio-economic benefits of new knowledge generated within research infrastructures from the scientific community's perspective. We use Earth Observation (EO) satellites as a case study. Constructing, operating, and exploiting cutting-edge EO infrastructures is generating a vast amount of knowledge ultimately embodied in scientific publications. Using bibliometric and machine learning techniques, we analyse 1,235 publications in 1998–2018 related to *Cosmo Skymed*, the EO satellites constellation of the Italian Space Agency. Thanks to these satellites, 2,377 authors from 160 institutions and 68 countries worldwide have contributed to various subjects in several scientific fields. By using the marginal social value method in a cost-benefit analysis perspective, we conservatively estimate the value of such publications, including their marginal cost and value of citations. This original and straightforward approach can be used to estimate the socio-economic value of scientific publications produced within any research infrastructure, including universities, in any field of study.

1. Introduction

Earth Observation (EO) is an increasingly relevant domain of the space sector consisting in the collection of a wide variety of chemical, biological and physical information about the planet earth, via remote sensing technologies (GEO, 2020). During the last years, EO infrastructures, including the number of satellites orbiting around the earth, have rapidly grown (SIA, 2018). They have become a highly strategic tool that provides a wide array of services and applications for governments, firms, scientists and citizens (PwC, 2016; Tassa, 2020). They contribute both to global issues such as climate change and air pollution monitoring and local ones such as precision farming, urbanisation, transport infrastructures management or citizen security, amongst others.

Globally, coalitions of countries and national space agencies are increasingly investing in EO infrastructures. However, these highly skilled labour and capital-intensive research infrastructures (RIs) are often controversial science policy investments. They need significant governments' financial efforts that ultimately involve taxpayers (Florio, 2019; Jacob and Hallonsten, 2012). Hence, before undertaking such investments, governments should carefully evaluate their convenience and return for society.

Research has already shown that the socio-economic benefits of public investments in large research infrastructures, including EO, are wide-ranging and encompass a multitude of different stakeholders, comprising firms, people, and the society in general (EC, 2014; Salter and Martin, 2001; Crawfard, 2016; Booz and Company 2014; Hof et al., 2012; PwC, 2019). Indeed, these benefits relate to diverse types of agents involved in different ways within the EO value chain, from the *upstream* sector, which includes the producers of EO satellites components to the so-called *downstream* industry, which embraces the intermediate and final users of the EO collected data.

While a discussion on the different current and potential payoffs of EO is out of the scope of this paper, we focus here on the estimation of the socio-economic benefits associated with a particular category of stakeholders: the scientific community. Projecting, building, operating and exploiting EO infrastructures typically require the involvement of a vast community of scientists, universities, research institutes and research & development (R&D) firms' departments. For scientists, directly and/or indirectly involved in projects such as the construction and launch of new satellites or the archiving, elaboration and use of data collected by satellites, one of the main benefits is the generation and/or

* Corresponding author.
E-mail address: davide.vurchio@uniba.it (D. Vurchio).

https://doi.org/10.1016/j.techfore.2022.121730
Received 23 March 2021; Received in revised form 11 March 2022; Accepted 29 April 2022
Available online 11 May 2022

acquisition of *new knowledge*. Thanks to the existence of EO infrastructures, scientists can access new information, technologies, methods and data that, most likely, would not be available otherwise. Such available information is then potentially used to generate an additional stream of knowledge, pushing the knowledge frontier forward.

The objective of this paper is to estimate the socio-economic benefits of EO infrastructure for the scientific community. Our main hypothesis is to distinguish between the value of knowledge *per se*, whose socio-economic impact is not quantifiable, and the value of a knowledge' output', such as a scientific publication, whose socio-economic impact can be quantified (Rosseau et al., 2020; Florio, 2019; Florio et al., 2016). Indeed, newly generated knowledge is usually incorporated and disseminated through scientific publications and used by the scientific community itself to create new knowledge outputs. Hence, in a social cost-benefit analysis perspective, our research question is: *What is the marginal social value of a statistical publication? What does the scientific community earn from the fact that an additional publication related to EO exists?*

We propose an original conceptual framework by looking at the socio-economic impact of a 'statistical' publication proxied by its marginal social value. This approach draws on the well-known cost-benefit analysis techniques to value a 'statistical life', which do not attempt to measure the intrinsic value of the human life rather determine how much individuals are willing to pay for a small reduction in the probability of death (Viscusi and Aldy, 2003).

We test this original conceptual framework by looking at the case of Cosmo Skymed (CSK- *COnstellation of small Satellites for Mediterranean basin Observation*), the constellation satellites of the Italian Space Agency (ASI).[1] Indeed, in Europe, both the European Space Agency (ESA) and ASI have largely financed the development of dedicated EO satellite constellations, respectively the Copernicus Programme and Cosmo Skymed. Italy is at the forefront of this sector; it has mastered a broad set of space technologies, from the production of launchers and satellites, flight control to data transmission and elaboration.[2]

In this conceptual and empirical exercise, we take advantage of insights generated by recent works addressed at estimating the social value of scientific publications (e.g., Rosseau et al., 2020), and lessons learned from the cost-benefits analysis of large research infrastructures (EC, 2014; Florio, 2019; Florio et al., 2016; Battistoni et al., 2016). The empirical application of this method is easy to implement; it also leads to conservative estimates and can be used to appraise the value of scientific publications produced by any research infrastructure, including universities, in any field of study.

This paper is organised as follows: in Section 2, we explain the conceptual framework; in Section 3, we provide general information on Cosmo Skymed, its properties and technical characteristics and an overview of related scientific publications and community of scientists. Lastly, in Section 4 we calculate the social value of such scientific publications while in Section 5 we draw our conclusion and outlook.

2. Estimating the benefits of Earth Observation (EO) for the scientific community: a conceptual framework

For centuries, knowledge, inventions, technology, scientific discovery have driven the progress of our society. Today is the 'information age' where knowledge has become the key asset for long-run socio-economic growth, besides capital and labour (Jaffe and Trajtenberg, 2002). In this scenario, large research infrastructures are crucial since they are facilities, resources and services that foster knowledge and innovation in different fields (EC, 2019). They are characterised by large investments in technological and human capital, and their primary goal is to produce scientific knowledge.

However, measuring in quantity and quality the socio-economic impact of knowledge generated by the existence of a research infrastructure, in a cost-benefit analysis perspective, is not always feasible. When this information is concretely used, for example, to undertake practical decisions, methods such as the Value of Information (VOI) allow quantifying the welfare change between a state in which a certain action is taken, based on currently available information, and a different state in which the same action is taken using improved information (Macauley, 2006; Gallo et al., 2018). In other cases, new knowledge is patented, and the socio-economic benefit can be estimated by looking at the marginal social value of the patent (Florio, 2019). However, scientific subjects often contribute to expanding understanding about certain issues (knowledge *per se*) and do not always lead to practical applications (although, of course, ultimately, each can contribute to the other). Indeed, it is not always possible to know in advance what will be the practical implications of acquiring a new piece of knowledge, what is its potential impact on the society and how this impact can be measured in monetary terms. For example, in the EO upstream sector, it is extremely challenging to determine in advance what could be the consequence of the invention of a new material needed to operate an EO satellite, how this new material can be re-used to create new products – or to improve performance of existing ones – and, in general, how such invention could impact our daily lives. In the same vein, in the EO downstream sector, it is not always possible to know *ex-ante* the socio-economic impact of acquiring new information on planet earth through remote sensing technologies. Hence, measuring the socio-economic impact of newly acquired knowledge is not always feasible, and any attempt would lead to uncertainty.

However, a feature of the contemporary production of scientific knowledge is the generation of a large number of scientific publications (Florio, 2019; Carrazza et al., 2016) (see for example Arxiv, 2019[3] and the growing number of scientific publications in different fields of science). Scientific publications include books, research papers, conference proceedings, amongst others. As explained in Rosseau et al. (2020) and Florio (2019), it is possible to estimate the value of a knowledge 'output' such as a scientific publication that embodies a piece of knowledge deriving, in this case, from the EO infrastructure. Indeed, large research infrastructures are publications factories (see, for example, Carrazza et al., 2016). By being involved (even indirectly) in an EO infrastructure, scientists can access and collect new information that can be further elaborated to produce a knowledge output ultimately published in the scientific domain. This knowledge output can, in turn, be used by the community of scientists to build upon, pushing the knowledge frontier forward.

Hence a critical point of this analysis is that we make a clear distinction between the value of knowledge *per se* which is unknown or not easily measurable from the value of a knowledge *output,* which is a scientific publication. In practice, we separate the value of publications from the value of scientific results, and we focus exclusively on the use-value from the scientists' perspective (Rosseau et al., 2020).

Besides the stream of publications as 'vehicles' of new knowledge, one could argue that another significant benefit for scientists is represented by the fact that, being involved in such breakthrough arrangements and publishing more, they can improve their skills, reputation and prestige (Salter and Martin, 2001). This benefit could be potentially estimated by looking at the increased opportunities and salary premiums that scientists can obtain in the job market after spending a period of work related to the EO infrastructure (Catalano et al., 2021). In this vein, Tuckman and Leahey (1975) estimate the economic return of publications for different cohorts of academic researchers. However, as explained in (Catalano et al., 2021), this effect is hard to measure within

[1] https://www.asi.it/en/earth-science/cosmo-skymed/

[2] https://www.iai.it/en/pubblicazioni/new-space-economy-opportunity-italy

[3] https://arxiv.org/help/stats/2019_by_area/index

research infrastructures, except for early career researchers, as the career advancement of established scientists depends on different other factors besides publications, for example, managerial qualities, networking capabilities and so on. Hence, in this paper, we focus exclusively on the benefit for the scientific community of new knowledge generation embodied in publications. At the same time, we neglect the effect that these publications can have on their career. We also neglect the effect that a greater number of publications can possibly have in attracting higher amounts of public funds.

2.1. The marginal social value of scientific publications

To estimate the socio-economic impact of scientific publications, we try to assess the marginal social value of a 'statistical' publication, that is the increased benefit, for the community of scientists, associated with the existence of an additional publication (Drèze and Stern, 1987; Boardman et al., 2017).

This evaluation may rely on different approaches, such as market prices or other methods based on individual preferences (see Rosseau et al., 2020 for a review). However, market prices are a poor proxy of the value of scientific publications as the publishing market is a highly distorted one (due to the situation of oligopoly,[4] presence of subsidies and other externalities, i.e., knowledge spillovers). Prices are often discriminated between individuals and institutions, and they do not provide accurate information on the actual preferences and 'use- values' of such publications from the point of view of the scientific community (Larivière et al., 2015; Rosseau et al., 2020).

In a cost-benefit analysis framework, the marginal social value can be estimated from a consumer or from a producer perspective. From a consumer perspective, we should consider the Willingness To Pay (WTP) of scientists for an additional publication. This would mean observing (with the method of *revealed preferences*) or asking (with the method of *stated preferences*), or with a combination of these two methods, how much an average scientist (the main consumer of scientific research) would be willing to pay for the availability of an additional publication (Carson et al., 1996; Breidert, 2007; Johnston et al., 2017). However, these methods cannot be very promising because usually, access to publication is rarely bought by scientists; they are freely available at physical or virtual libraries or traditionally paid by third parties (e.g., academic institutions). Additionally, the WTP based on stated preferences relies on surveys where subjective opinions depend on socio-economic and cultural conditions, which can vary substantially amongst an international and varied scientific community of scientists such as those involved in research infrastructures. Launching a survey, as in Rosseau et al., 2020 can be an effective method when looking at one delimited institution, although costly and time-consuming. These authors look at the WTP of scientists to publish, which is conceptually different from the WTP to acquire an additional publication where scientists are considered the primary consumers of scientific knowledge. They collect 315 responses at the University La Sapienza of Rome, with a response rate of 15%. However, when looking at research infrastructures, scientists are dispersed amongst several institutions worldwide; hence interviewing a representative sample of the reference population becomes challenging. Additionally, another important issue would be getting scientists to tell the truth. The so-called 'hypothetical bias' is the difference between monetary values people say they would be willing to pay in a hypothetical scenario, with what an individual might actually pay in the real setting (Florio and Giffoni, 2020). Policy consequentiality techniques are strictly connected to the truth-telling issue and are often used for improving the realism of the survey by conveying to respondents the idea that their choice will have consequences (Florio and Giffoni, 2020). Hence ex-ante hypothetical bias and ex-post mitigation measures are essential to estimate the WTP more robustly but are not easily applicable to the publishing market (Arrow et al., 1993; Johnston et al., 2017; Rosseau et al., 2020).

From a producer perspective, a more promising method consists in the estimation of the marginal cost of a publication which refers to the increased cost associated with producing an additional publication (Drèze and Stern, 1987; Boardman et al., 2017). In practice, because generally in the case of research infrastructures, consumers and producers of publications match (the scientific community), we assume that the implicit price the scientific community is willing to pay for an additional publication is at least the cost of producing it. Indeed, a specific feature of research infrastructures is that the demand for new knowledge is driven by the scientific community, which is also the producer of such knowledge.

In the last few years, the cost of scientific publishing has attracted the attention of a bunch of scholars, especially from the publishers' perspective (Bergstrom and Bergstrom, 2004; Bergstrom, 2001; SQW, 2004; Van Noorden, 2013; Grossmann and Brembs, 2021). Here, we start from a general cost function of scientific publications subjected to a peer-review process that includes fixed and variable costs of two main tasks i) researching and concretely writing a publishable content (*research costs*), ii) transforming the publishable content in a publication (*publishing costs*). Building on the cost and business model in scientific research publishing (SQW, 2004), and on a social cost-benefit analysis framework, we exploit and refine the original model proposed by Florio (2019) so that the typical social cost function of scientific publications ($Total\ Cost_{scpub}$) can be expressed as follows:

$$Total\ Cost_{scpub} = f\left(K_{res}, L_{res}, L'_{res}, OP_{res}, \right) + f\left(K_{pub}, L_{pub}, L'_{pub}, OP_{pub}\right) + EXT_{res} + EXT_{pub} \quad (1)$$

Where K_{res} is the capital cost of scientists producing publishable content (entirely ascribable to their research institute of affiliation). It includes, for example, the cost of laboratory equipment, computing resources, software, databases, and any other fixed asset needed for research, amongst others. L_{res} is the labour cost of scientists expressed in terms of wage rates and time devoted to research. L'_{res} is the administrative and technical labour that support their research activity, OP_{res} are other operative costs ascribable to the research institutes of affiliation (e.g. electricity, chemical or biological materials, costs of trips for research, and conferences, amongst others).

In the second term of the function, K_{pub} is the capital costs of publishers (e.g. computing resources software licences, and any other fixed asset needed for creating online or printed publications); L_{pub} is the labour cost of publishers (e.g. work of editors and referees in the peer review process); L_{pub}' is the administrative and technical labour of the entire publication process (e.g. proofreaders, copywriters etc.); OP_{pub} are other operative costs of publishers (e.g. paper distribution costs for printed journals, maintenance of the electronic system for online journals) (see SQW, 2004 for a review of main cost components of publications).

Lastly, in a social cost-benefit analysis perspective, EXT_{res} are externalities imputable to any effect of the research process (e.g., use of toxic reagents, ethical issues, etc.) while EXT_{pub} are externalities related to the publication process (e.g. CO2 emission for printing and distributing pieces of paper).

In the computation of the marginal cost of publications, by taking the first derivative of the function cost with respect to the number of publications, we can assume that all capital costs, administrative and technical labour for the researchers and publishers are fixed and do not change as the number of publications changes (SQW, 2004). Hence $\frac{\Delta K_{res}}{\Delta pub} = 0$; $\frac{\Delta L'_{res}}{\Delta pub} = 0$; $\frac{\Delta K_{pub}}{\Delta pub} = 0$; $\frac{\Delta L'_{pub}}{\Delta pub} = 0$.

Therefore the marginal cost is mainly a function of the direct workforce associated with the publication, including the cost of scientists, editors and reviewers where $\frac{\Delta L_{res}}{\Delta pub} > 0$ and $\frac{\Delta L_{pub}}{\Delta pub} > 0$ plus a minor

[4] The publishing market is characterized by a limited number of big players such as Reed-Elsevier, Wiley Blackwell, Springer, and Taylor & Francis (Larivière et al., 2015)

portion of operative expenses which can vary with publications. Indeed we assume that a large majority of operative costs – both for the researchers and publishers - do not change as the number of publications changes (e.g. electricity, material for experiments, paper distribution costs for printed journals); hence we assume $\frac{\Delta OP_{res}}{\Delta pub} = 0$ and $\frac{\Delta OP_{pub}}{\Delta pub} = 0$, although this is not always the case (e.g. cost for conferences and trips can increase as the number of publications increases). For simplicity, we also neglect the role of externalities imputable to any effect of the research and publication process (e.g. CO2 emissions caused by repeated experiments or by the physical distribution of papers). In the absence of microdata at the publication level, we are not able to identify and evaluate the operative costs and research externalities that are specific to different research projects and publication outlets. Still, we can simply acknowledge that, for some publications, these might be higher than zero. By neglecting these aspects, the estimated marginal social cost publications – on which we rely to compute the marginal social benefit – will be lower than the actual one, confirming our estimates' conservative nature.

Previous works simply estimate the marginal cost by adopting the opportunity cost of scientists' time approach. (EC, 2014; Florio, 2019; Rosseau et al., 2020). The idea is that when scientists spend time on a research project related to EO, which may then translate into an additional publication, they have an opportunity cost for not working on an alternative project. Hence, a monetary value can be attached to a scientific publication by estimating its marginal cost, which depends on the salary received by the scientists (in our function, we also consider the opportunity cost of editors' and reviewers' time) and the time dedicated to that publication. While Rosseau et al. (2020) in their work also attempt to apply this method by interviewing professors at the University La Sapienza of Rome, we further develop and advance this approach by exploiting available secondary data at the publication level. While the availability of microdata at the publication level could provide a more precise estimation of different items of the cost function (e.g. operative costs of a specific research project, time devoted to research, etc.) our approach is more efficient, especially when the focus is not a single institution but a diverse community of scientists from all over the world.

Hence, the marginal production cost *MPC* of a representative paper published at time *t* can be computed as the first derivative of the total cost function:

$$MPC_t = MPC_{res_t} + MPC_{pub_t} = \left(\frac{w_{res_t} * h_{res_t}}{y_{res_t}}\right) + \left(\frac{w_{pub_t} * h_{pub_t}}{y_{pub_t}}\right) \tag{2}$$

where w_{res_t} is the average gross annual wage of scientists employed in research, h_{res_t} is the average share of time researchers spend in producing their publications and y_{res_t} is their average scientific productivity, which is a function of the total number of publications (considering authors' both EO and non-EO publications) produced in the year *t*.[5] w_{pub_t} is the average gross annual wage of editors and reviewers employed in the peer-review process, h_{pub_t} is the share of the time they employ for peer reviewing and y_{pub_t} is the yearly average number of papers they peer review.[6] As mentioned, additional variable costs, including paper distribution costs with decreasing importance in recent years, and other externalities are not determinable without microdata at the publication level and thus approximated to zero.

Moreover, a new publication has a higher value for the scientific community when it is used to generate new knowledge outputs. Hence to obtain the marginal social value of a 'statistical' publication, we need to augment its marginal cost with the value of its influence. The latter can be a function of the number of citations received. Using citations as a proxy for the influence of a publication has several drawbacks (Harzing, 2010; Waltman et al., 2013). For example, citations vary significantly across different fields (Harzing, 2010). Clarivate[7] in 2019 reveals that publications related to space science receive a number of citations that is, on average, more than three times higher than the number of citations received by publications in other sectors benefitting from the existence of EO infrastructures, such as computer science. However, citations are commonly considered a reflection of the impact that a particular piece of publication has generated, particularly in the field of science (Harzing, 2010; Bourke and Butler, 1996).

In this context, in our framework, we should consider at least the first round of publications citing the initial publications directly related to the EO infrastructure.[8] To assign a monetary value to citations, we shall again consider the opportunity cost of time employed by a scientist to read and cite someone else' publication. Thus, the shadow price of citation is estimated using the opportunity cost of time employed by a scientist to decide to cite a publication which can vary from few minutes to hours and depends on several factors (Florio, 2019). Since this information is difficult to obtain, we compute the value of citation of a representative paper ($ValueCIT$), at year t, by multiplying the marginal cost of a publication by a coefficient (α) equal to the share of time that researchers, on average, devote to reading activities. Then we divide the results by the average number of references reported in the citing papers ($avRef_t$) roughly assuming that each of these references has contributed equally to the production of the publication (Florio, 2019) .

$$ValueCIT_t = \alpha * MPC_t / avRef_t \tag{3}$$

Following this reasoning, the marginal socio-economic value (MSV_t) of a publication is expressed by the value of the scientists, editors and reviewers' time (MPC_t) augmented by the value of its impact in the literature ($ValueCIT_t$)(Florio et al., 2016; Florio, 2019).

$$MSV_t = MPC_t + ValueCIT_t \tag{4}$$

As the last step to estimate the socio-economic value of EO publications, we compute the present value at time, $PV_{EO,\tau}$, as:

$$PV_{EO,\tau} = \sum_{t=1}^{T} EO_t * MPC_t \Big/ (1+SDR)^{t-\tau} + \sum_{t=1}^{T} CIT_t * ValueCIT_t \Big/ (1+SDR)^{t-\tau} \tag{5}$$

Where EO_t is the total number of publications related to EO produced in the considered period, CIT_t is the total number of citations received by EO_t and *SDR* is the usual social discount rate adopted in cost-benefit analysis (EC, 2014).

The strength of this approach is that it is easy to apply with publicly available data. It is also very conservative as usually the social value proxied by the marginal cost is lower than those calculated with WTP (Boardman et al., 2017). Additionally, as highlighted by Romer (1990), by using market wages of scientists, we underestimate the social value of scientific research as market wages are commonly lower than shadow wages (the latter are the social opportunity costs of labour and should be

[5] When authors work on different publications, this aspect is captured by their annual productivity.

[6] In our theoretical setting, we use the yearly average salary of researchers and the share of time devoted to research/peer reviewing activities to assign a value to the representative (marginal) article. However, the reader can easily adapt such setting to the case in which the information is already available in monetary units, by substituting one or both the members in equation (2) with the monetary value of the marginal cost of publication for researchers and/or for publishers. Indeed, Florio (2019) builds his framework on hourly unit of measures, not needing to derive the marginal value from annual salaries.

[7] https://clarivate.com/webofsciencegroup/solutions/journal-citation-reports/

[8] Other subsequent round of citations could be considered estimating a sort of epidemiology of the publication as explained in Carrazza et al. (2016). Something that is out of the scope of the paper.

used in a cost-benefit analysis framework to take into account of labour market distortions, although rarely available). Indeed generally, scientists are underpaid in the wage market as the role of knowledge spillovers and other positive externalities is not taken into consideration (Romer, 1990). We also neglect to simplify additional possible operative costs and externalities, further shrinking the estimation of the social value.

In the next section, we introduce a practical application of this framework to the case of Cosmo Skymed the satellite constellation of ASI.

3. The case of Cosmo Skymed satellite constellation

Entirely developed by ASI in cooperation with the Italian Ministry of Defence, Cosmo Skymed constellation satellites is based on advanced remote sensing technology and has both military and civilian applications. It is composed of four identical satellites[9] orbiting the earth at 619 km of altitude and with a 97-minute orbital period allowing a high temporal resolution, and its system can carry out up to 450 shots of the earth's surface per day, equal to 1800 radar images, every 24 h.[10] Since its first launch in 2007, the constellation has contributed to post-earthquake, hurricanes and cyclones activities, monitoring deforestation in Amazonia and the ice caps shrinking in the polar regions. It has also supported sustainable farming, surveillance of UNESCO sites and oil spill monitoring, amongst others.[11]

Earth Observation represents the main programme of ASI. *Fig. 1* provides a general overview of the budget allocation in 2013 (OECD, 2014), where the total national allocation of ASI in Earth Observation was about EUR 123 million. Besides investment in national programmes, in 2013 ASI also invested about EUR 108 million in the ESA' Earth observation Programmes, reaching a total of about EUR 131 million. Additionally, by looking in detail at ASI's investments in EO programs, Cosmo Skymed plays a key role: in 2013, the programme received more than EUR 87 million,[12] considering investments both in industry and research.

ASI gives paramount importance to research activities. Since its establishment, it has carried out its activities by collaborating with many firms and research centres active in the space sector. Table 1 shows the distribution of contracts with universities, research institutes and institutions by programmes, as reported in the *Archimede* dataset.[13] The information refers to the period 1996–2018 and provides information on EUR 614 million.

About half of this amount refers to research activities related to universe exploration and observation (48.5%), while EO and Cosmo-Skymed Program activities have received less than 23% of the invested resources (about EUR 140 million). According to the description reported in the database, 66% of EO contracts value is represented by Cosmo Skymed contracts while, by excluding the Ministry of Defence that is both the main ASI's partner in Cosmo Skymed programmes and amongst the main funders of EO activities,[14] contracts with research institutes sum up to EUR 51.6 million, differentiated as reported in Table 2.[15]

The majority of contracts reported in the dataset has been signed with public institutions: 52.2% of the contracts, corresponding to 97 out of 186 contracts, refer to non-university partners such as the European Space Research Institute (ESRIN),[16] Italian National Institute for Nuclear Physics (INFN), Italian National Institute for Geophysics and Volcanology (INGV) and National Research Council (CNR). Such a slight majority corresponds, in terms of contracts value, to a 65.5%: this reveals how partnerships with public institutes is characterised by contracts of the higher amount with respect to partnerships with academic researchers. amongst the main university partners, we find the Department of Electrical, Electronic and Information Engineering of the University of Bologna, the Department of Physics of University of Rome Tor Vergata and the Department of Physics of the University of Rome Sapienza.

3.1. *Peer-reviewed scientific publications related to Cosmo Skymed*

In this section, we provide an overview of the scientific publications related to Cosmo Skymed. We explore the number and types of publications available on *Scopus*, which is the largest database of peer-reviewed literature, including scientific journals, books and conference proceedings.[17] In this way we can explore the contribution of the constellation satellites in generating reliable knowledge (Bornmann and Marx, 2013).

By using the Scopus search Application Programming Interface (API)[18] we extrapolated and analysed information on peer-reviewed documents related to Cosmo Skymed, published during the period 1998–2018. Our search query allowed us to extract all peer-reviewed documents and related information, containing the words "Cosmo Skymed" in their title, keywords or abstract for a total of 1235 documents. 36% of the extracted documents are classified as journal articles, 62% are conference proceedings, while the remaining 2% includes books and reviews.[19]

Fig. 2 shows how the number of publications experienced a rapid increase since the year 2007, when the first and the second satellite of Cosmo Skymed constellation were launched, respectively in June and December 2007, reaching a peak of 164 publications in 2012. In the following years, data suggest that after the last launch in November 2010, the interest for the technologies brought by Cosmo Skymed satellites has been stationary and, in some cases, decreased.

On average, each paper has been cited by other researchers 10.6 times for a total of 13,088 direct citations, with an increasing trend until 2019. The median citation, taking into account possible distortions caused by the presence of outliers (e.g., papers that have been particularly influential and therefore frequently cited) is three (Table 3).

According to the Scopus classification (Table 4), slightly less than one-third of publications are in the earth and planetary sciences, while computer science and engineering are, respectively, the second and the third subject areas, represented, in total, by 40% of the documents. Relatively less represented are publications in physics and astronomy, mathematics and material sciences. Additionally, it is remarkable how publications that could widely benefit from the collection of EO data on planet earth (e.g., social, environmental or agricultural sciences) are

[9] In this work we focus on first generation satellites. We neglect second generation satellites (first launch in 2019, out of our period of analysis).
[10] https://www.researchitaly.it/en/projects/cosmo-skymed-earth-satellite-observation-dual-system/.
[11] https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/cosmo-skymed
[12] http://ww2.gazzettaamministrativa.it/opencms/export/sites/default/_gazzetta_amministrativa/amministrazione_trasparente/_agenzie_enti_stato/_agenzia_spaziale_italiana/130_bila/010_bil_pre_con/2014/Documenti_1410872262201/1410872263557_consuntivo_2013.pdf
[13] *Archimede* is a repository containing information on the contracts that ASI has signed since 1989 with firms, universities, and research institutes.
[14] 94% of contracts amount refers to contracts signed with such institutions.
[15] Since the aim of this work is to understand the social value of a publication relative to Cosmo Skymed, we avoid considerations relative to the military use of the Italian constellation of satellites.
[16] https://www.esa.int/About_Us/ESRIN
[17] https://www.scopus.com/home.uri
[18] https://dev.elsevier.com/documentation/ScopusSearchAPI.wadl
[19] The query has been performed in October 2020.

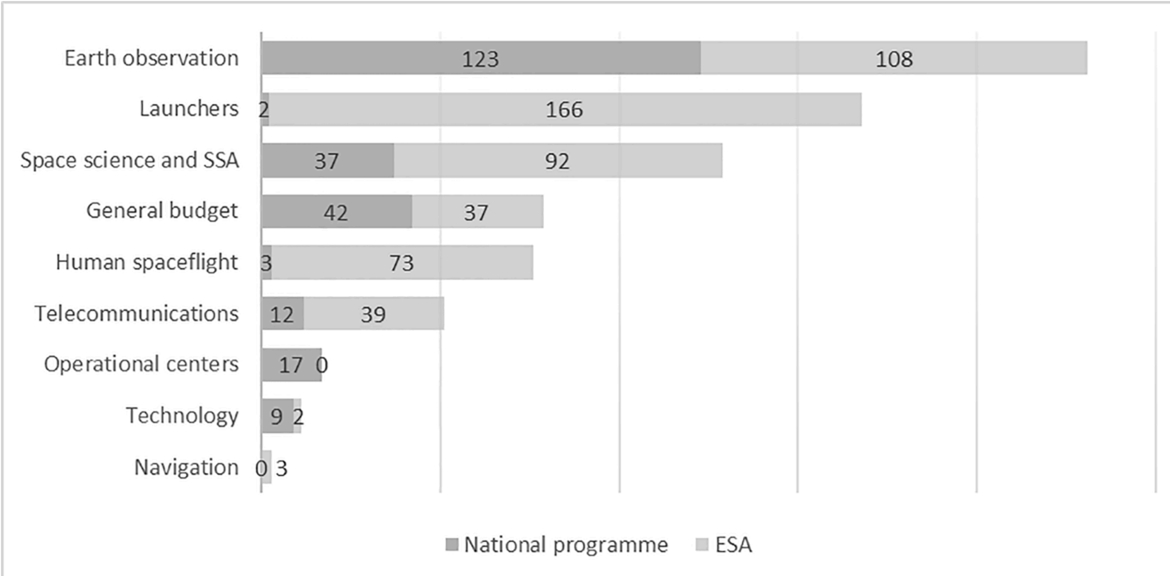

Fig. 1. Italian space agency's budget by main programmes in 2013 (millions of euros) **Source:** OECD (2014), calculations based on ASI, 2014.

Table 1
Italian Space Agency funds to universities and research centres.

Area	(%)
Universe exploration and observation	48%
Earth observation and Cosmo Skymed	23%
Human space flight and microgravity	11%
Telecommunication and applications	6%
Not specified	5%
Technologies and Technological Transfer	3%
Strategic direction	2%
Bases management	1%
Launchers and Space surveillance	1%

Source: Own elaboration on Archimede dataset.

Table 2
Italian Space Agency's contracts by type of partner (excluding the Ministry of Defence) (1996–2018).

Contractor	N. of contracts	N.of contracts (%)	Value (EUR)	Value (%)
Public institutes	97	52.15	33,814,520	65.57
Universities	89	47.85	17,759,390	34.43

Source: Own elaboration on Archimede dataset.

underrepresented, suggesting that the use of Cosmo Skymed in the downstream segment of the value chain is still unexploited. Although Cosmo Skymed has gained much attention for its cutting-edge technological features, the exploitation of data for civil use is still limited by its military applications and data policy. Additionally, EO data are complex and not user-friendly. They need transformation, manipulation, and integration with other types of data to provide meaningful information. Hence, the engagement of institutions, including universities and other research centres, in using Cosmo Skymed data can be particularly restrained due to several factors, including the access policy, the lack of expertise and inertia to innovation (Tassa, 2020) (see NEREUS, 2016 for a review of the main obstacles for data exploitation). Conversely, other satellites such as Sentinels by ESA or Landsat[20] by the National Aeronautics and Space Administration (NASA) provide free-open access and user-friendly services to maximise their exploitation, although the use of data is still limited (Florio and Morretta, 2021)[21].

To classify more precisely the contents of the extracted documents, we adopted a generative model based on unsupervised machine learning able to extract topics from a set of documents and to provide an explanation on the similarity of individual parts of the documents: the Latent Dirichlet Allocation (LDA) method (Blei et al., 2003).[22] With this methodology, the title and abstract of each document is considered as a set of words that, combined together, form one or more subsets of latent topics characterised by a particular distribution of words. The idea behind is that when a paper is related to a certain topic, particular words are expected to appear more frequently. Hence, the main topics captured by this technique are clusters of similar words. The model allows the classification of the main topic of a given document by providing the probability of being in a specific topic class in case the document discusses multiple topics. The algorithm helped us to identify 7 topics (Fig. 3). After applying the LDA method, we performed a manual check of 50% of the papers of the sample to see whether the topic automatically assigned to the document was reflecting the real content.

The first cluster of documents, composed by 35.6% of publications, is mainly related to *Landslide, subsidence and ground deformation observation*. The second cluster of topics (14.7%) refers to *Cosmo Skymed capabilities and operative modes* (with several works focusing on the design of second-generation satellites), while the third cluster relates to *Flood analysis, water and resource monitoring* (13.8%). The fourth cluster of documents represented by 12.6% is related to *Crop monitoring, cryosphere observation and wind detection*. The fifth cluster of documents focuses on *Processing, interferometry and over time scene analysis* and includes 9.2% of publications. The least represented topics are *Imagery analysis, sensor accuracy and resolution* and *Building detection, high-resolution imagery and damage maps* including, respectively, 7.3% and 6.9% of documents.

Focusing on the list of authors, we find that 2377 unique names appear on the title page of the selected documents. These authors belong to 160 different institutions and are located in 68 countries around the world.

Fig. 4 and Table 5 present the distribution of authors according to their institutions and country of affiliation. Publications related to

[20] https://www.usgs.gov/core-science-systems/nli/landsat

[21] See for example https://www.copernicus.eu/en/copernicus-services

[22] https://radimrehurek.com/gensim/about.html

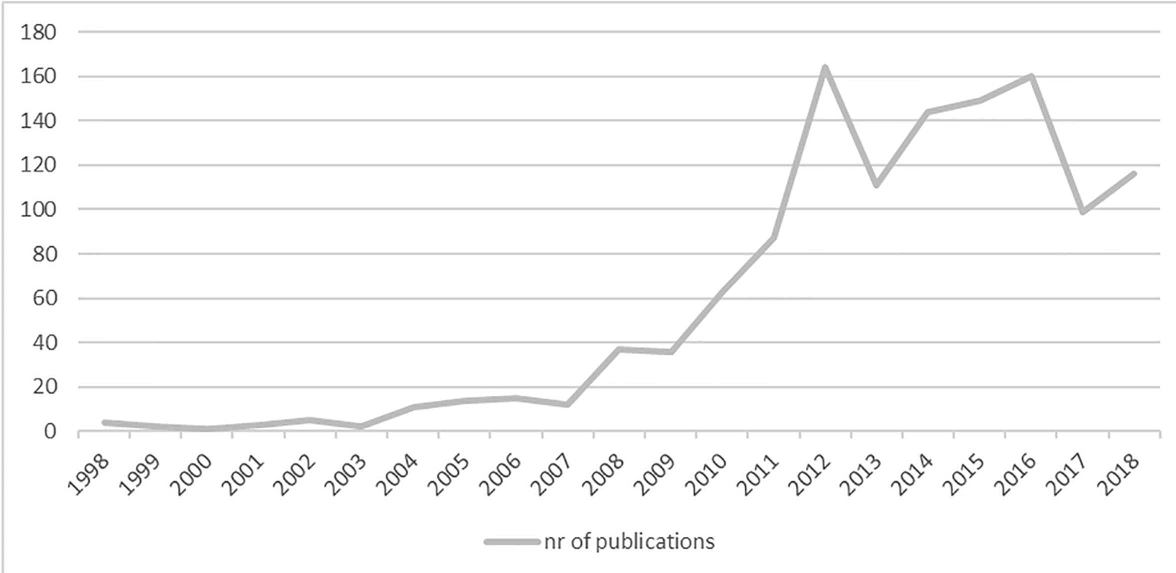

Fig. 2. Number of peer-reviewed scientific publications (1998–2018) **Source:** Own elaboration on Scopus.

Table 3
Publications related to Cosmo Skymed (1998–2018).

Total number of papers	1235
Total number of authors	2377
Total number of citations	13,088
Citation Average	10.6
Citation Median	3

Source: Own elaboration on Scopus.

Table 4
Cosmo Skymed publications by subject area.

Subject	N. of publications (%)
Earth and Planetary Sciences	31.2%
Computer Science	23.3%
Engineering	16.7%
Physics and Astronomy	9.8%
Mathematics	5.3%
Materials Science	4.7%
Environmental Science	3.2%
Social Sciences	2.5%
Agricultural and Biological Sciences	1.5%
Other	1.9%

Source: Own elaboration on Scopus.

Cosmo Skymed are mainly produced by authors affiliated with Italian institutions (869 authors, 53.6% of the total), while the first alternative is represented by authors with Chinese affiliations (111 authors, 6.8%).

This territorial bias, which may also reflect the difficulty of using Cosmo Skymed data and other information outside Italian borders, can also be found by looking at the list of the main institutions to which the authors are affiliated which are mostly Italian. ASI is the main institute of affiliation since it is represented by a total of 220 authors (9.3% of the total). All other institutes are universities or public research bodies (such as the National Research Council and the National Institute of Geophysics and Volcanology) except for the Italian-French firm Thales Alenia Space Italia, the main European producer of satellites, however subject to partial public control.

4. Estimating the marginal social value of publications: an application to Cosmo Skymed

As previously mentioned, an important step for the estimation of the marginal cost of Cosmo Skymed publications is to estimate authors' productivity. In doing so, we extrapolated the full amount of documents, not necessarily related to Cosmo Skymed, written during the period 1998–2018 by the 2377 authors. In total, these authors have written 69,652 documents on different topics.

The literature on authors' productivity indices usually assigns full credit for a joint paper to each co-author (e.g., Lehmann et al., 2006 and Perry and Reny, 2016). Our baseline estimates are computed by assigning full credit of a publication to each co-author, meaning that, on average, each author has produced 4 documents per year regardless of the number of co-authors. As an alternative, we divide credits uniformly amongst the yearly average number of co-authors (e.g. Hirsch, 2007; Shen and Barabasi, 2014) thus ignoring that each co-author might have contributed differently to the publication of the documents (Table 6).

Intermediate methods to estimate authors' productivity are also contemplated and would deserve further and *ad hoc* analyses (e.g. Flores-Szwagrzak and Treibich, 2016), although they are out of the scope of the paper.

As the second step of our analysis, we shall determine the cost of writing a document which can be estimated by looking at the salary received by scientists to write their research. Since we do not have information at the scientist level, we retrieved researchers' salary information from the European Commission (EC, 2007). This study collects information on 6190 researchers across Europe and other non-European countries, including China and the United States, with a response success rate of more than 50%. 72% of interviewed researchers conduct research activities at higher education institutes, 20% at governmental institutions, while 8% work in the business sector. 40% of the interviewed researchers work in engineering science, information science, environment and geoscience and physics, 27% in life science, 25% in social and human science while the remaining percentage (9%) in mathematics and physics. All contacted researchers declared to devote at least 50% of their time to research (EC, 2007).

The strength of this study, compared to similar surveys (e.g., Altbach et al., 2012; Rumbley et al., 2008), is that it provides salary information for a sample of heterogeneous countries and researchers, which is consistent with our reference sample, and it adopts a transparent and robust methodology. In the absence of information on shadow wages,

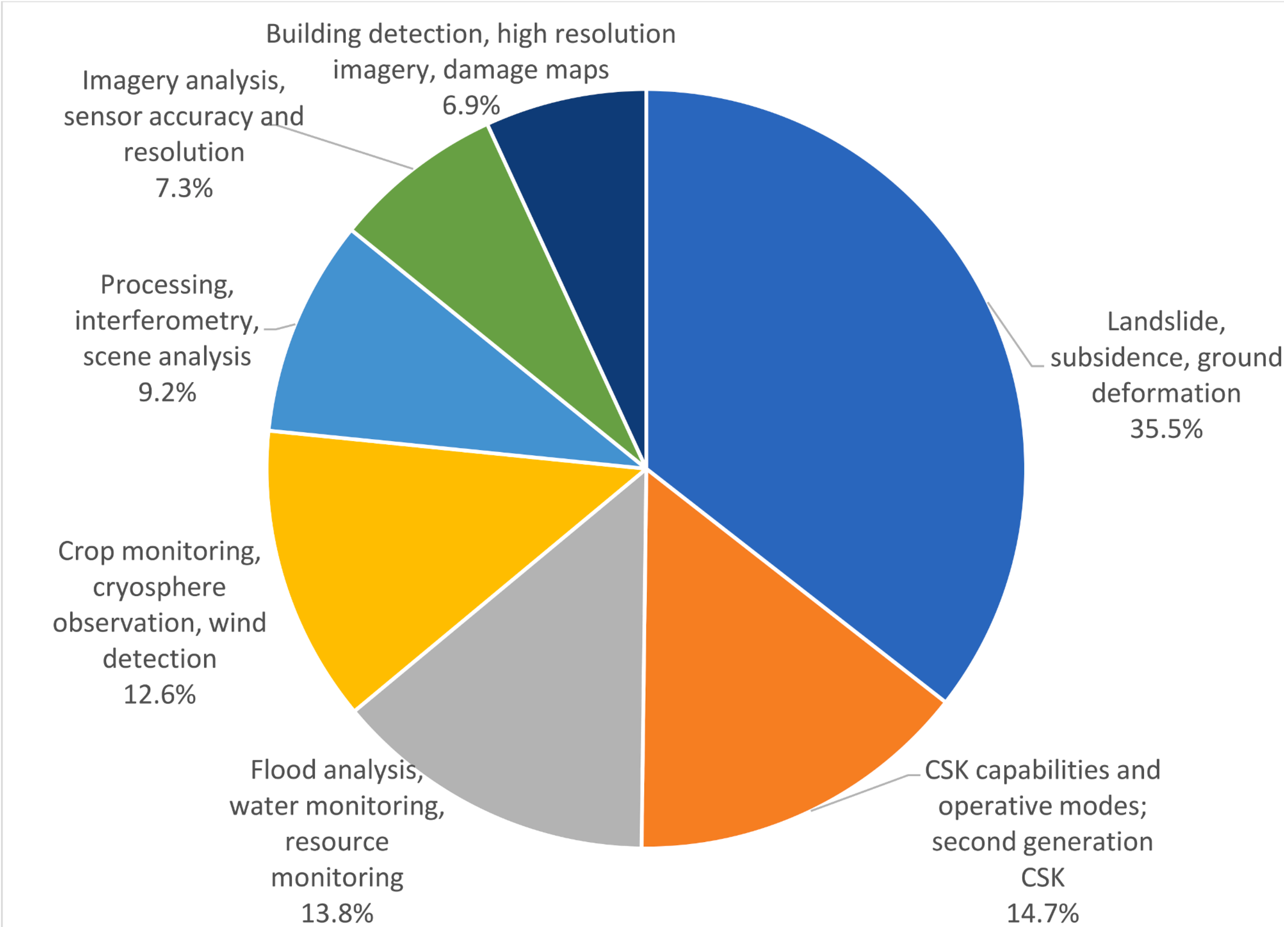

Fig. 3. Cosmo-Skymed topics from Latent Dirichlet Allocation method **Source:** Own elaboration on Scopus.

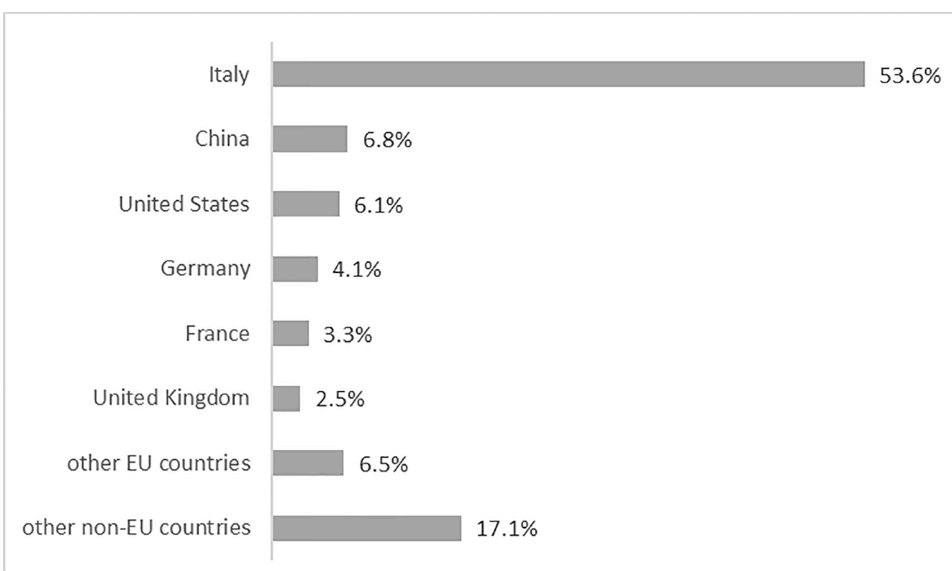

Fig. 4. Number of authors by country of affiliation **Source:** Own elaboration on Scopus.

Table 5
Number of authors by top affiliation.

Institution of affiliation	%
Agenzia Spaziale Italiana	9.3%
National Research Council	6.9%
Università degli Studi di Roma La Sapienza	4.0%
National Research Council, Institute for Electromagnetic Sensing of the Environment	3.4%
Università degli Studi di Napoli Federico II	2.8%
Thales Alenia Space Italia	2.8%
Istituto Nazionale Di Geofisica E Vulcanologia, Rome	2.7%
Parthenope University of Naples	2.7%
Università degli Studi di Roma Tor Vergata	2.0%
Istituto Di Studi Sui Sistemi Intelligenti Per L'automazione, Bari	1.8%
Politecnico di Bari	1.8%
Università degli Studi di Firenze	1.6%
Università degli Studi di Genova	1.6%
Istituto Di Fisica Applicata Nello Carrara	1.6%
Telespazio S.p.A.	1.6%
Deutsches Zentrum fur Luft- Und Raumfahrt	1.3%
Jet Propulsion Laboratory	1.3%
Università degli Studi di Trento	1.3%
Chinese Academy of Sciences	1.2%
California Institute of Technology	1.2%

Source: Own elaboration on Scopus.

Table 6
Average authors' productivity (including publications out of EO domain).

Year	Full credit attribution method	Partial credit attribution method
1998	2.9	1.1
1999	3.0	1.0
2000	3.2	1.0
2001	3.1	1.0
2002	3.5	1.0
2003	3.4	1.1
2004	3.7	1.2
2005	4.1	1.4
2006	4.1	1.3
2007	4.0	1.2
2008	4.2	1.2
2009	4.2	1.2
2010	4.3	1.2
2011	4.2	1.1
2012	4.5	1.1
2013	4.5	1.1
2014	4.6	1.1
2015	4.8	1.1
2016	4.6	1.0
2017	4.4	1.0
2018	5.0	1.1
Average	4.0	1.1

Source: Own elaboration on Scopus.

the inclusion of salaries of researchers in governmental institutions and private markets, which are argued to be more efficient compared to the highly subsidised academic sector,[23] is able to proxy the opportunity cost of time of researchers more adequately. Another important strength of this survey is that it provides information on Purchasing Power Standard (PPS) to take into consideration the different cost of living in each country as calculated by Eurostat.[24] Additionally, salary data on

Table 7
Average gross salary in research by country (baseline year: 2006).

Country	Remuneration average in EUR	Remuneration average in PPS
Austria	62,406	60,530
Belgium	58,462	55,998
Cyprus	45,039	50,549
Czech Republic	19,620	36,950
Denmark	61,355	43,669
Estonia	11,748	21,053
Finland	44,635	36,646
France	50,879	47,550
Germany	56,132	53,358
Greece	25,685	30,835
Hungary	15,812	27,692
Ireland	60,727	49,654
Italy	36,201	34,120
Latvia	10,488	21,580
Lithuania	13,851	29,660
Luxembourg	63,865	56,268
Malta	28,078	40,342
Netherlands	59,103	56,721
Poland	11,659	21,591
Portugal	29,001	33,334
Slovakia	9178	18,282
Slovenia	27,756	37,970
Spain	34,908	38,873
Sweden	56,053	47,143
United Kingdom	56,048	52,776
Bulgaria	3556	9770
Croatia	16,671	27,063
Romania	6286	13,489
EU-28 average	34,829	37,624
Iceland	50,803	33,801
Israel	42,552	59,580
Norway	58,997	41,813
Switzerland	82,725	59,902
Turkey	16,249	26,250
Australia	64,150	62,342
China	3150	13,755
India	9177	45,207
Japan	68,872	61,991
United States	60,156	62,793
Worldwide	37,685	40,024

Source: Own elaborations based on EC (2007).

which we base our estimations are provided for the year 2006, which is approximatively in the middle of the period under scrutiny. The possibility to consider a year, not at the beginning or the end of the period under scrutiny helps to mitigate the problem of using the same salary information, although we adopt constant prices for each year.

The first column of Table 7 shows the total yearly salary, expressed in EUR, for researchers considered in the EC (2007) 's analysis. It includes the net salary received, employers' charges (e.g., social security contribution, pension funds), employee contribution to social security, holiday pay and personal income tax. The second column shows the same salary at PPS.

High differences between countries often reduce when considering the cost of living. However, the EU25 average (37,624 EUR) is far below the US average (62,793 EUR).

Taking into account the country distribution percentages of authors of publications related to Cosmo Skymed (Fig. 4), and applying the average salaries by country, we find – for 2006 – an average salary in our sample of EUR 37,475 (PPS 37,537). As the information on salaries we collected is available only for 2006, by adopting country-specific Gross Domestic Product deflators provided by the World Bank[25] and setting 2006 as the base year, we have been able to compute salary values for the entire period 1998–2018. By multiplying the average annual gross

[23] While the marginal cost can be a good proxy for the value of a publication, academia is not a well-functioning competitive labor market as it is heavily subsidized and regulated. Additionally, as stated in Rosseau et al. (2020) transaction costs limit international mobility hence the assumption of a well-functioning market is unlikely to hold.

[24] The original Purchasing Power Parities are standardized using the EU25 average as baseline. This artificial unit (currency) is named PPS (Purchasing Power Standard). The baseline year adopted is 2006 that is in the middle of our time span and allows us to have data for most of the affiliation countries (some of which reported in Figure 4).

[25] https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS

Table 8
Marginal cost of a statistical publication for researchers, by year (MPC_{res_t}).

year	Full credit attribution method EUR	Full credit attribution method PPS	Partial credit attribution method EUR	Partial credit attribution method PPS
1998	5375	5335	14,648	14,539
1999	5306	5269	16,112	16,001
2000	5119	5088	16,341	16,241
2001	5322	5294	16,978	16,892
2002	4898	4878	17,067	16,998
2003	5066	5051	16,279	16,231
2004	4840	4836	15,517	15,503
2005	4466	4467	13,486	13,491
2006	4607	4614	14,701	14,725
2007	4807	4828	15,979	16,048
2008	4659	4690	16,304	16,413
2009	4820	4855	16,740	16,863
2010	4669	4716	16,541	16,709
2011	4912	4977	18,221	18,463
2012	4660	4730	18,616	18,896
2013	4780	4858	19,674	19,997
2014	4707	4787	20,012	20,356
2015	4560	4640	19,897	20,248
2016	4764	4854	21,681	22,090
2017	5021	5128	22,543	23,021
2018	4574	4682	21,452	21,961
Average	4854	4885	17,561	17,699

Source: Own elaboration on Scopus and EC (2007).

salary of a researcher by the share of time dedicated to research[26] and dividing by authors' productivity, we calculated the first term in Eq. (2), i.e., the opportunity marginal cost for researchers related to Cosmo Skymed. In table 8, we report values (both in EUR and in PPS) for both attribution methods. In the first case, publications are full-credited to each co-author (full credit). In the second case, each co-author is credited a share of publication (partial credit).[27] Please note that the MPC_{res_t} component attributed to the opportunity cost of researchers' time is inversely proportional to the average yearly productivity. By crediting each co-author with (only) a share of publication, the total number of publications released by all the authors is lower than in the full credit attribution method. This leads to lower productivity that, subsequently, increases the marginal cost value.

Further details are provided by looking at the MPC_{res} by country. Fig. 5 shows that publications with the highest marginal research cost (full publication attribution method) are those written by Swiss researchers (about EUR 16,330 per publication), Austrian researchers (about EUR 14,560 per publication) and Luxembourg's researchers (about EUR 13,350 per publication). Considering the values in PPS, the highest marginal costs of research are those associated with publications by Israeli, Czech and Maltese researchers. The lowest values, both in EUR and PPS are recorded for publications written by Chinese scientists. Italian researchers, who represent over 53% of Cosmo Skymed authors, have a relatively low average marginal cost: this might lead to lower benefits associated with Cosmo Skymed publications. By using the partial publication attribution method, higher costs are observed in Malta, Denmark and Check Republic.

Although the marginal cost of a publication for researchers represents the major component of the marginal cost of producing a scientific publication, to effectively compute the latter as in Eq. (2), we need to consider the marginal cost of the publishing activity, which include both the editorial and peer-review process costs. Several studies attempt at estimating the cost of publishers for several types of journals. For instance, based on costs estimation from the literature and from surveys, SQW (2004) appraise an article first-copy cost between 250 and 2000 US dollars depending on the quality of the journal. Dubini et al. (2012), by interviewing several international publishers, find an average cost of 250 US dollars for peer-review management and 170 to 400 US dollars for formatting, editing, typesetting and metadata. Van Noorden (2013) provided mixed evidence. By relying on interviews, he estimates the average total cost of publication in about 3700 US dollars for *Proceedings of the National Academy of Sciences* and much higher values for *Nature*. Grossmann and Brembs (2021) provide a step-by-step calculation of the costs associated with publishing primary research articles, from submission, through peer-review, to publication, indexing and archiving. The authors find that publication costs range from less than 200 US dollars per article in modern, large scale publishing platforms using post-publication peer-review,[28] to about 1000 US dollars per article in prestigious journals with rejection rates exceeding 90%. The publication costs for a representative scholarly article are quantified in about 400 US dollars. In a scenario regarding professionally run journals and peer-reviewers providing volunteer labour (traditional commercial journal) the cost of publication is equal to 723.16 US dollars. In a scenario taking into account scholarly journals operating with a minimal budget, not paying their editors and using institutional servers with free, open-source Open Journal System handling submission, the cost of publication is equal to 237.35 US dollars. In our analysis, in the absence of information about reviewers and editors' productivity and time devoted to publication-related activities, we use the first amount as a proxy of the cost of publication for journal articles and books while we adopt the second amount as a proxy of the cost for conference proceedings. 723.16 and 237.35 US dollars are the values we assign to 2018. By mean of US GDP deflator values, euro-dollar yearly average exchange rate and euro PPS-dollar yearly average exchange rate, we obtain the values reported in Table 9.[29]

As previously mentioned, to obtain the marginal social value of a 'statistical' publication, we need to augment its marginal cost with the value of its influence which is a function of the number of citations received (see Eq. (3)). For simplicity, we consider only the first round of publications citing the initial publications directly related to the EO infrastructure without considering additional streams of subsequent citations (Carrazza et al., 2016).[30] To consider the opportunity cost of time employed by a scientist to read and cite someone else's publication, we refer to the time academic scholars devote, on average each year, to read scientific articles and books. We follow STM (2018) and Tenopir et al. (2012), who estimate this time in 448 h, and assuming an average yearly amount of time spent in research activities equal to 1000,[31] we identify a α parameter equal to about 0.45. To estimate Eq. (3) we also need the average number of references reported in the papers citing Cosmo Skymed papers that ranges from a minimum of 16 in 1998 to a maximum of 73 in 2002.

Considering the average marginal cost of a publication times the number of Cosmo Skymed publications[32] plus the average value of a citation multiplied by the number of citations (see Eq. (5)) and by evaluating all values at time $\tau = 2020$, we obtain the results reported in

[26] In line with EC (2007), we assume that researchers devote half of their time to their publications (h_{res}=0.5). h_{res} is constant during the entire period.

[27] Share equal to the reciprocal of the number of coauthors.

[28] For further information see: https://www.mededpublish.org/What-is-post-publication-peer-review

[29] Deflator values are retrieved from https://data.worldbank.org/indicator/NY.GDP.DEFL.ZS, nominal exchange rate are retrieved from https://www.macrotrends.net/2548/euro-dollar-exchange-rate-historical-chart and EUR PPS - US dollar exchange rate are retrieved from https://ec.europa.eu/eurostat/web/purchasing-power-parities/data/main-tables. Missing values in 1998 and 1999 are substituted by exchange rates in 2000.

[30] Neglecting the impact of following streams of citations would return a value that represents the "lower bound" in a conservative way.

[31] We assume an average worker employed for 40 hours *per* 50 weeks. Given the parameter h assumed equal to 0.5, the total amount of time devoted to research is equal to 1,000.

[32] Considering a full publication with the full credit attribution method and only a share of publication, considering co-authors, with the partial credit attribution method.

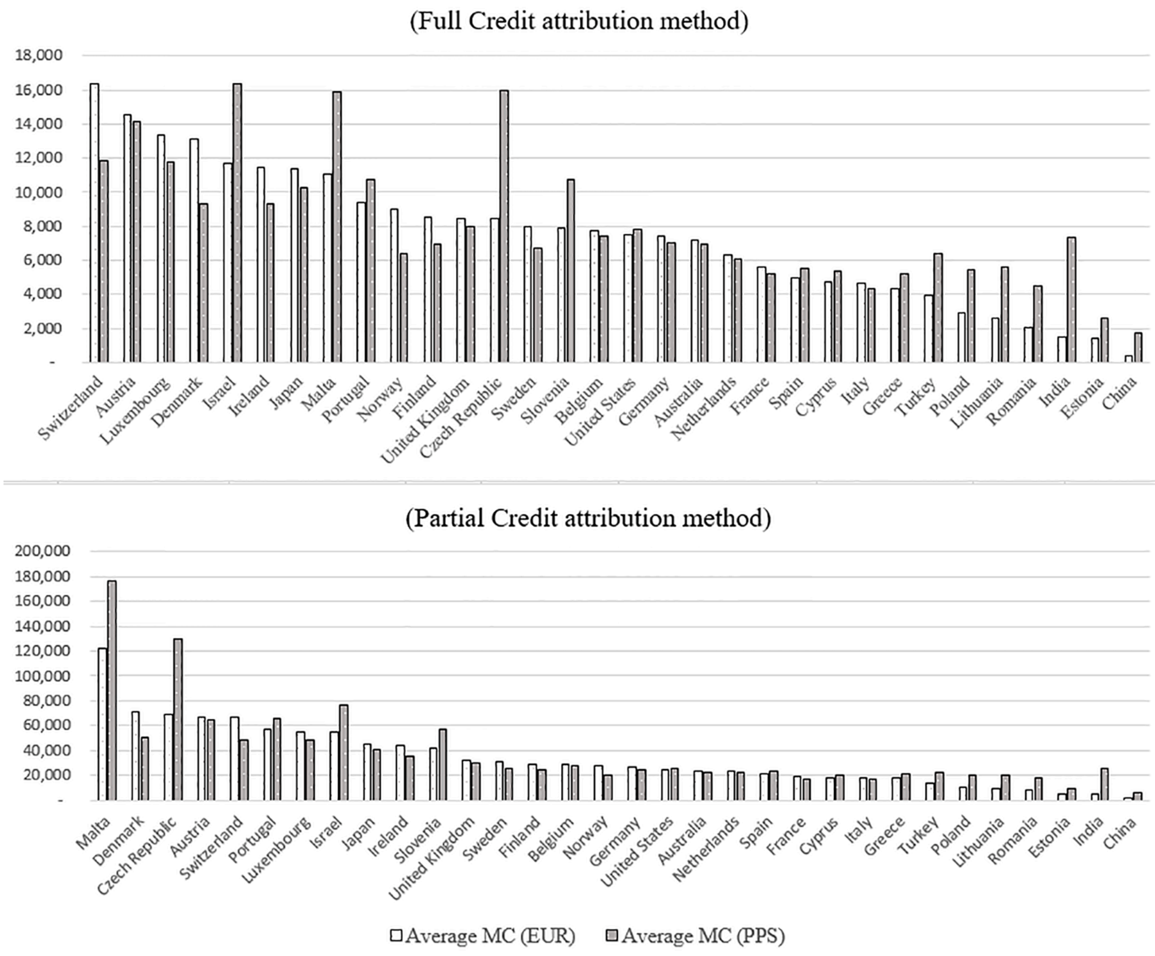

Fig. 5. Marginal cost of a statistical publication for researchers, by country (MPC_{res_t}) **Source:** Own elaboration on Scopus and EC (2007).

Table 9
Marginal cost of a statistical publication for publishers, by year (MPC_{pub_t}).

year	EUR Conference proceedings	Journal articles and books	PPS Conference proceedings	Journal articles and books
1998	175.77	535.55	139.56	425.21
1999	178.31	543.28	141.57	431.35
2000	182.30	555.43	144.74	440.99
2001	190.44	580.22	147.40	449.09
2002	183.27	558.38	148.34	451.95
2003	156.94	478.15	152.33	464.13
2004	146.86	447.47	156.07	475.52
2005	151.44	461.41	160.18	488.04
2006	153.55	467.82	160.00	487.48
2007	145.01	441.82	163.10	496.95
2008	137.77	419.77	163.24	497.36
2009	146.81	447.32	162.03	493.69
2010	155.23	472.94	163.51	498.18
2011	151.63	461.98	164.40	500.88
2012	166.52	507.34	166.26	506.56
2013	164.34	500.72	165.02	502.80
2014	167.45	510.19	167.48	510.27
2015	202.79	617.85	169.94	517.79
2016	205.00	624.60	166.11	506.12
2017	205.20	625.21	167.18	509.38
2018	201.14	612.85	169.23	515.61
Average	169.89	517.63	158.94	484.25

Source: Own elaboration on Grossmann and Brembs (2021).

table 10. According to our analysis, the present social value of publications related to Cosmo Skymed ranges, very conservatively, between EUR 32.4 and 37.4 million at PPS..[33] This estimate does not include any social use of the knowledge embodied in the papers. This different topic deserves further research, and that is beyond the scope of our empirical analysis.

5. Concluding remarks

The present work proposes and develops an original conceptual framework to estimate the socio-economic benefits related to new knowledge generation, from the strict perspective of the scientific community. As a case study, we consider Earth Observation. Projecting, building and operating EO infrastructures, as many research infrastructures, require the involvement of a vast community of scientists, universities, research institutes and R&D firms' departments.

In this work, we adopt a social cost-benefit analysis perspective, and we distinguish between the value of knowledge *per se* and the value of a knowledge output. Indeed, newly generated knowledge is commonly incorporated and disseminated through scientific publications and used by the scientific community to push the knowledge frontier forward. Hence, we estimate the value of scientific publications, produced in the

[33] This analysis has been carried out during the month of October 2020. All values have been capitalized to 2020 since the power $t-\tau$, is negative. The social discount rate adopted is 3 % (EC, 2014).

Table 10
Present social value of Cosmo Skymed publications.

	Full credit attribution method		Partial credit attribution method	
	EUR	PPS	EUR	PPS
PV of publications	36,317,985.00	36,837,794.00	29,728,117.00	30,156,104.00
PV of citations	519,118.59	523,315.41	2177,751.50	2216,245.75
PV publications and citations	**36,837,104.00**	**37,361,108.00**	**31,905,868.00**	**32,372,350.00**

Source: *authors' elaborations.* The value of publications and citations is evaluated at time τ=2020.

period 1998–2018, related to Cosmo Skymed, the EO satellites constellation of ASI. By exploiting bibliometric and machine learning techniques, we assess a marginal social value of EO scientific publications obtained by summing the value of the research activity behind each publication, the value of publication-related activities and the value of citations. On average, and according to the method by which publications are credited to authors and the type of publication, the value of researcher activity is between EUR PPS 4885 and 17,699; the value of publisher activity is between EUR PPS 159 and 484 and the value of a single citation is between EUR PPS 52 and 192 with the total benefit of Cosmo Skymed for the scientific community ranging between EUR PPS 32.4 and 37.4 million. The marginal social value represents the average value the scientific community would earn from the existence of an additional publication related to Cosmo Skymed. In other words, it is the value to access new pieces of information related to EO (i.e., cutting edge technologies, analytical methods and data on planet earth) that, most likely, would not be available otherwise. The empirical application of this method also leads to conservative estimates of the investigated benefit as we focus exclusively on peer-reviewed literature, while many publications are also found within the 'grey literature'. We also use scientists' market wages as shadow wages of scientists are not available and the former are usually lower than the latter (Romer, 1990). Furthermore, estimation methods based on marginal cost generally lead to more conservative estimates of the marginal social value compared to the WTP methods and we only consider the value of the first round of publications without considering additional streams of citations. We also neglect unknown operative costs and externalities. Moreover, we neglect the value of the scientific results, which is uncertain and can be reasonably assumed higher than zero. We must also remember that benefits for scientists are just a portion of the vast benefits generated by EO infrastructures, which include a large community of stakeholders such as firms, people and the society in general.

From this analysis, we can state that EO literature is rapidly growing. The application to Cosmo Skymed suggests a significant social value of this literature, although restricted data use policy may restrain the scientific potential as compared to other non-military open-access data programmes such as Copernicus (ESA) or Landsat (NASA). Several technical, legal, organisational or communication barriers are still hindering the full exploitation of information that can be extrapolated from Cosmo Skymed, an issue that deserves further research to improve ASI data governance, particularly within the downstream value chain.

The strength of this original approach to estimate the socio-economic value of scientific research, compared to stated or revealed preferences methods, is that it is easy to implement and can be used to estimate the value of scientific publications produced by any research infrastructure or university in any field of study. Secondary data analysis has the advantage of being low-cost and easy to implement compared to collecting primary data through surveys. Although the primary collection of data could provide additional information on different social costs of publications, including externalities and deliver more precise estimations, as mentioned, such method is not convenient for large research infrastructures characterised by a wide variety of scientists located worldwide. Further research is, however, needed to fine-tune the application of this methodology for instance by including subsequent rounds of citations to estimate – more precisely – the value of the influence of a publication in the literature or by accounting for externalities (e.g. emissions related to research projects in different disciplines or for printing publications). In addition, future exercises based on sensitivity analysis could also be helpful to assess the impact of assumptions which can have a profound impact on the estimation of the final benefit such as the scientists' time devoted to research activity which in our study, following EC (2007), has been set to 50% of working time. Indeed this hypothesis is not necessarily realistic for different career stages or diverse private and public institutions within different countries.

Funding

This paper has been developed within the framework of the project *"Cost-Benefit Analysis of public policy in the space industry"* carried out by the University of Milan and funded by the Italian Space Agency (ASI).

CRediT authorship contribution statement

Valentina Morretta: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Davide Vurchio:** Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Writing – review & editing. **Stefano Carrazza:** Formal analysis, Data curation.

Acknowledgements

We wish to thank Masimo Florio, Lorenzo Zirulia, the editor and the two reviewers for their helpful suggestions, which have been crucial to refocusing, enriching and strengthening the content of this paper. Earlier versions of this paper have been presented to the *Milan European Economic Workshop* 2020 and the *Society of Cost-Benefit Analysis* 2020.

References

Altbach, Philip G., Pacheco, Ivan F., 2012. Paying the professoriate: A global Comparison of Compensation and Contracts. Routledge.
Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA panel on contingent valuation. Fed Regist 58, 4601–4614.
Battistoni, G., Genco, M., Marsilio, M., Pancotti, C., Rossi, S., Vignetti, S., 2016. Cost–benefit analysis of applied research infrastructure. Evidence from health care. Technol. Forecast. Soc. Change 112, 79–91.
Bergstrom, C.T., Bergstrom, T.C., 2004. The costs and benefits of library site licenses to academic journals. Proc. Natl. Acad. Sci. 101 (3), 897–902.
Bergstrom, T.C., 2001. Free labour for costly journals? J. Econ. Perspect. 15 (4), 183–198.
Blei, D.M., Ng, A.Y., Jordan, M.I., 2003. Latent dirichlet allocation. J. Mach. Learn. Res. 3 (Jan), 993–1022.
Boardman, A.E., Greenberg, D.H., Vining, A.R., Weimer, D.L., 2017. Cost-benefit analysis: Concepts and Practice. Cambridge University Press.
Booz & Company (2014): Evaluation Of Socio-Economic Impacts From Space Activities In The Eu, Prepared for: The European Commission Enterprise And Industry Directorate-General, March 2014.
Bornmann, L., Marx, W., 2014. How should the societal impact of research be generated and measured? A proposal for a simple and practicable approach to allow interdisciplinary comparisons. Scientometrics 98 (1), 211–219.
Bourke, P., Butler, L., 1996. Publication types, citation rates and evaluation. Scientometrics 37 (3), 473–494.
Breidert, C., 2007. Estimation of willingness-to-pay: Theory, measurement, Application. Springer Science & Business Media.
Carrazza, S., Ferrara, A., Salini, S., 2016. Research infrastructures in the LHC era: a scientometric approach. Technol. Forecast. Soc. Change 112, 121–133.

Carson, R.T., Flores, N.E., Martin, K.M., & Wright, J.L. (1996). Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. Land Econ., 80–99.

Catalano, G., Giffoni, F., & Morretta, V. (2021). Human and social capital accumulation within research infrastructures: the case of CERN. Ann. Public Coop. Econ.

Crawford, I.A., 2016. The long-term scientific benefits of a space economy. Space Policy 37, 58–61.

Drèze, J., Stern, N., 1987. The theory of cost-benefit analysis. Handbook of Public Economics. Elsevier, pp. 909–989. Vol. 2.

Dubini, P., Di Trani, F. and Micheli, M.R. (2012), PEER Economic Report. Retrieved from http://www.peerproject.eu/fileadmin/media/reports/PEER_Economics_Report.pdf.

EC - European Commission (2007). Remuneration of Researchers in the public and private sector, April 2007 https://cdn1.euraxess.org/sites/default/files/policy_library/final_report.pdf.

EC - European Commission (2014). Guide to Cost-Benefit Analysis of Investment Projects Economic appraisal Tool For Cohesion Policy 2014-2020 https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/cba_guide.pdf.

EC. European Commission (2019) European research infrastructure 'what research infrastructures are, what the commission is doing, strategy areas, funding and news' https://ec.europa.eu/info/research-and-innovation/strategy/european-research-infrastructures_en.

Flores-Szwagrzak, K., Treibich, R., 2015. Co-authorship and the measurement of individual productivity. University of Southern Denmark Discussion Paper on. Bus. Econ. 17.

Florio, M., 2019. Investing in Science: Social Cost-Benefit Analysis of Research Infrastructures. MIT Press.

Florio, M., Giffoni, F., 2020. A contingent valuation experiment about future particle accelerators at CERN. PLoS ONE 15 (3), e0229885.

Florio, M., Morretta, V., 2021. Earth observation and economic studies: a cross-fertilisation perspective. Space Policy 57, 101429.

Florio, M., Forte, S., Sirtori, E., 2016. Forecasting the socio-economic impact of the Large Hadron Collider: a cost–benefit analysis to 2025 and beyond. Technol. Forecast. Soc. Change 112, 38–53.

Gallo, J., Bernknopf, R., Lev, S., 2018. Measuring the socio-economic value of data and information product derived from earth observations and other geospatial products. In: Kruse, JB, Crompvoets, J, Pearlman, F (Eds.), GEOValue: the Socio-Economic Value of Geospatial Information. Taylor & Francis Group, London, pp. 33–48.

GEO (2020) "What is earth observation" https://www.earthobservations.org/g_faq.html.

Grossmann, A., Brembs, B., 2021. Current market rates for scholarly publishing services. F1000Res 10 (20), 20.

Harzing, A.W., 2010. The Publish or Perish Book: Your guide to Effective and Responsible Citation Analysis. Tarma Software Research, Melbourne.

Hirsch, J., 2007. Does the h index have predictive power? Proc. Natl. Acad. Sci. 104 (49), 19193–19198.

Hof, B., Koopmans, C., Lieshout, R., Wokke, F., 2012. Design of a Methodology to Evaluate the Direct and Indirect Economic and Social Benefits of Public Investments in Space. SEO rapport 42.

Jacob, M., Hallonsten, O., 2012. The persistence of big science and megascience in research and innovation policy. Sci. Public Policy 39.4 (2012), 411–415.

Jaffe, A.B., Trajtenberg, M., 2002. Patents, citations, and innovations: A window On the Knowledge Economy. MIT press.

Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., et al., 2017. Contemporary guidance for stated preference studies. J. Assoc. Environ. Resour. Econ.; 4, 319–405. https://doi.org/10.1086/691697.

Larivière, V., Haustein, S., Mongeon, P., 2015. The oligopoly of academic publishers in the digital era. PLoS ONE 10 (6), e0127502.

Lehmann, S., A. J. and Lautrum, B. (2006). Measures for measures. Nature 444 (7122), 1003–1004.

Macauley, M.K., 2006. The value of information: measuring the contribution of space-derived earth science data to resource management. Space Policy 22 (4), 274–282.

NEREUS (2016), Improving Copernicus take up among Local and Regional Authorities (LRAs) via dedicated thematic workshops ESA Contract No. 4000110850/14/I-BG http://www.nereus-regions.ovh/wp-content/uploads/2017/11/Analysis.pdf.

OECD, 2014. The Space Economy At a Glance 2014. OECD Publishing.

Perry, M., Reny, P.J., 2016. How to Count Citations If You Must. Am. Econ. Rev., Am. Econ. Assoc. 106 (9), 2722–2741. , pagesSeptember.

PwC, (2016) Study to examine the socio-economic impact of Copernicus in the EU Report on the Copernicus downstream sector and user benefits October 2016.

PWC, (2019): Copernicus Market Report, February 2019. Prepared for the EC.

Romer, P.M., 1990. Endogenous technological change. J. Polit. Econ. 98 (5), S71–S102.

Rousseau, S., Catalano, G., Daraio, C., 2020. Can we estimate a monetary value of scientific publications? Res. Policy 50 (1), 104116.

Rumbley, L., Pacheco, I.F., Altbach, P.G., 2008. International Comparison of Academic salaries: An exploratory Study. Boston College Center for International Higher Education, Chestnut Hill, MA.

Salter, A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. Res. Policy 30 (3), 509–532.

Shen, H., Barabasi, A., 2014. Collective credit allocation in science. Proc. Natl. Acad. Sci. 111 (34), 12325–12330.

SIA -Satellite Industry Association (2018): State of the satellite industry report 2018 available at https://www.satelliteconfers.org/wp-content/uploads/2018/12/Therese-Jones-Satellite-Industry-Association-.pdf.

SQW (2004) Costs and Business Models in Scientific Research Publishing A report Commissioned By the Wellcome Trust, Cambridgeshire https://wellcome.org/sites/default/files/wtd003184_0.pdf.

STM (2018). The STM Report An Overview of Scientific and Scholarly Publishing. Retrieved from https://www.stm-assoc.org/2018_10_04_STM_Report_2018.pdf.

Tassa, A., 2020. The socio-economic value of satellite earth observations: huge, yet to be measured. J. Econ. Policy Reform 23 (1), 34–48.

Tenopir, C., Volentine, R., King, D.W., 2012. Scholarly Reading and the Value of Academic Library Collections: results of a study in six UK universities. Insights 25 (2), 130–149. https://doi.org/10.1629/2048-7754.25.2.130.

Tuckman, H.P., Leahey, J., 1975. What is an article worth? J. Polit. Econ. 83 (5), 951–967.

Van Noorden, R., 2013. Open access: the true cost of science publishing. Nature. 2013 495 (7442), 426–429.

Viscusi, W.K., Aldy, J.E., 2003. The value of a statistical life: a critical review of market estimates throughout the world. J. Risk Uncertain. 27 (1), 5–76.

Waltman, L., van Eck, N.J., Wouters, P., 2013. Counting publications and citations: is more always better? J. Informetr. 7 (3), 635–641.

Valentina Morretta is a post-doc researcher in Public Economics at the Department of Economics, Management and Quantitative Methods of the University of Milan. She has extensive experience in project cycle management, research and analysis in socio-economic disciplines, including the evaluation of the socio-economic impact of innovation policies and tools for local and regional development. She has worked in the private and no-profit sector and in academia in different countries, including Ireland, Italy, Kosovo and the UK. Her main interests focus on the evaluation of innovation policies, cost-benefit analysis, regional and local socio-economic development.

Davide Vurchio is a research fellow in Economic Policy at the Department of Economics and Finance of the University of Bari "Aldo Moro". After getting his PhD in Economics at the University of Bari in 2016 he has worked as post-doc research fellow at the Department of Economics of the University of Roma Tre and at the Department of Economics, Management and quantitative Methods of the University of Milan. Davide Vurchio has also worked as research consultant for the CRES Rossi-Doria, the United Nations Industrial Development Organization, Cerpem and Ismeri Europa, contributing to the realization of policy reports.

Stefano Carrazza is Assistant Professor in Physics at the Department of Physics of the University of Milan. He has been a graduate student at École Normale Supérieure de Lyon (ENS), a PhD student at the University of Milan, and a CERN fellow. Mostly focused on computational models and methods applied to science, Carrazza gathered experience in quantum optics at the Institut Lumière Matière in Lyon, cavity quantum electrodynamics at the Laboratoire Kastler-Brossel (ENS) in Paris with the Nobel Price Serge Haroche, data analysis and software development for heavy-ion physics in ALICE at CERN, theoretical physics at CERN and UNIMI, and quantum computing scientific collaboration at the Quantum Research Centre.